

The Census in the USA and Germany: It's all about sampling

Susan Rose-Ackerman

2018-12-09T10:37:41

In Germany disputes over the 2011 census have finally come to an end in the [‘census judgment’ by the Bundesverfassungsgericht](#). Its opinion rejected claims by Hamburg and Berlin that the shift towards sampling in the last nationwide census was unconstitutional. In the United States of America, in turn, disputes over the 2020 Census questionnaire are currently before the federal courts that raise related issues of the accuracy of the count. Commerce Secretary Wilbur Ross ordered the Bureau of the Census to include a question on citizenship status that is now subject to judicial challenge. In the first case to reach the trial stage, the US Southern District Court in New York has completed oral hearings. Its holding is expected to generate appeals, and, indeed, the Supreme Court has already agreed to hear one challenge. The cases are occurring within a tight time frame; the final text of the questionnaire must be settled by early summer.

In contrast to the German Grundgesetz, the U.S. constitution in [Article 1, Section 2, Clause 3](#) requires a census every ten years to determine representation in the House of Representatives and the Electoral College. Accurate population data are the underlying goal of the Constitution’s decennial census requirement. We argue that in the twenty-first century accuracy requires modern statistical techniques, including sampling and adequate pre-testing of questions.

The decennial census is one of the Constitution’s few substantive mandates. The count is of “persons”, not adults or citizens or voters. Each state’s representation in the House and in the Electoral College depends on the accuracy of the census — as does the distribution of federal funds under many programs. In 1901, a federal Court found in [United States v. Moriarty](#) that the census need not be limited to a headcount of the population, and it stressed that the gathering of other statistics, if ‘necessary and proper’ is permitted.

In the service of accuracy, statistical theory and practice have developed massively since 1868 when Section 2 of the Fourteenth Amendment removed the 3/5 weight on slaves in the original text and required a count of “whole persons”. Applying the “original” understanding of statistics in 1789 or even after the Civil War is inconsistent with the original and present-day goal of an accurate count.

The U.S. Census Act accepts the value of sampling and states that the Bureau “shall” use sampling everywhere other than for the purposes of reapportionment ([13 U.S. Code §195 – use of sampling](#)). The Supreme Court in [Dept. of Commerce v. U.S. House of Representatives](#) read the statute as disallowing sampling for the census but [did not reach the constitutional issue](#). The Supreme Court permitted, in [Utah v. Evans of 2002](#), a technique called “hot induction” to estimate missing census data because it did not involve a survey. Thus, the Supreme Court is not hostile to

statistics per se, but has, so far, not reached the constitution issue. Justice Stevens' dissent, however, comes close:

Since it is perfectly clear that the use of sampling will make the census more accurate than an admittedly futile attempt to count every individual by personal inception, interview or written interrogatory, the proposed method is a legitimate means of making the "actual enumeration" that the Constitution commands.

Sampling techniques are sophisticated and widely used in business, academic, and government applications. The Census Bureau uses statistical sampling to estimate key measures such as the unemployment rate, which has been [based on sampling since 1937](#). Sometimes sampling is a cost-saving measure, but in other contexts it is more accurate than counting one-by-one. Yet, direct counts have a certain intuitive appeal. You point to a small pile of apples and ask your friend how many apples are in the pile. Your friend guesses "eight". You count the apples and discover that there are actually nine. Clearly, the count is more accurate than your friend's guess. But the superiority of a direct count evaporates as the number counted increases and if the subjects have minds of their own. People are not apples. If you take a census of human beings, their personal characteristics may affect their answers and their willingness to be counted. Even if the head counters claim that the state will not use census data to impose individualized costs or benefits, some may refuse to participate. This refusal could arise, for example, from a concern for personal privacy or from an unwillingness to reveal negative information.

The American Statistical Association, the American Sociological Association, and Population Association of America filed an [amicus curiae brief](#) in the New York case. It critiques the rushed way in which the question was added to the census questionnaire, which did not follow accepted professional norms. Those arguments, however, are short-term concerns that the Bureau could correct before the next census in 2030. More important is the accuracy of head counts compared with statistical sampling as ways to gather reliable demographic information. The professional associations claim that the citizenship question is unnecessary because the Bureau's American Community Survey already provides estimates of citizenship status broken down by geographical areas. The Bureau uses statistical techniques and sample surveys to increase the accuracy of the demographic material gathered in the census. If sampling is at least as accurate and if the citizenship question would reduce the accuracy of the head count, then sampling should be used to estimate the mix of citizens and noncitizens in the population, and omitted from the census questionnaire.

Following the case law, the plaintiffs in the district court case did not argue for anything other than an accurate direct count. Yet, statistical sampling can improve the accuracy of the census—as either a complement to or a substitute for a head count, especially in the face of a citizenship question. Because an accurate population count is the overriding aim of the census clauses, it is unconstitutional to outlaw sampling as a general matter. The ban on sampling will be especially troublesome if a citizenship question is added. If the federal courts continue to accept the ban, then the Constitution forbids inclusion of a citizenship question in a census that allocates House seats and Electoral College votes to the states because

it reduces accuracy. Alternatively, if the question remains, the Census Bureau must rely on sampling to produce an accurate head count if it is to fulfill its constitutional mandate to provide accurate population data.

Germany faced similar choices of technique when carrying out the 2011 census. Public resistance to providing personal data to the government has posed particularly acute problems. The [widespread public protests and boycotts of the census](#) in 1987, which arose to a lesser degree in the [2011 census](#), were one of the reasons why no head count census had been conducted in over 20 years. Lacking a constitutional mandate for periodic counts, successive governments delayed the collection of new data. However, faced with out-of-date information and under pressure from the European Union, the 2011 census aimed to [gather data that would be comparable with data from other EU Member States](#). However, recalling the earlier protests, the Statistical Office sought to achieve an accurate count while minimizing data gathering from individuals. As explained by the BVerfG in its 2018 '[Census case](#)', the office proposed to rely on official registration documents that cover about 90% of the population, supplemented with household surveys in jurisdictions where it expected the undercount to be high.

The city-states of Hamburg and Berlin objected to this mixed method because it promised to assign them lower population totals than the previous method, disadvantaging them in the allocation of social welfare funds. They did not challenge the use of statistics per se, but only the particular method employed by the government. The Federal Constitutional Court held in favor of the federal government, deferring to the methodology used by the Statistics Office as representing the best expertise available that also limited interference with the fundamental right to "informational self-determination" [para. 286]. The Court stated that

The current state of statistical science cannot establish a clear superiority of the full census method compared to a register-based survey. [...] In comparison to the full census, the register-based method [...] can be an additional cause for faults; but it also provides for the benefits of a greater accuracy [para. 285].

The German court held that as long as the experts certified that the government's techniques promised a count at least as accurate as a head count, the methodology would satisfy the German constitution. It recognized the popular appeal of a direct head count stating that:

the full census, as a traditional way to collect data, has also shaped the ideas of constituents with regard to the accuracy of the population count. In any case, one cannot demand more from a new mathematical-statistical procedure than what a full census can provide [para. 280].

So long as there was widespread agreement among the experts that the proposed methods would be at least as accurate as a head count, those methods should be available for use by the government.

Lacking a constitutional text, the German court deferred to statistical expertise. In the U.S., given a constitutional text, the mandate for accuracy is at least as strong as in Germany. The purpose of a decennial census, as understood by the drafters and by constitutional lawyers and citizens is a count that minimizes errors in the count while recognizing that 100% accuracy is impossible. Given that concession to reality, a statute that outlaws the most accurate methods of counting heads should be held unconstitutional. The meaning of the clause should be updated by U. S. courts to accord with state-of-the-art knowledge. Thus, if a citizenship question reduces the accuracy of the head count, the data must be supplemented with statistical sampling and surveys to counteract those errors. Conversely, if omitting the question will increase accuracy, that finding should be sufficient to outlaw its inclusion under the Constitution.

A version of this blog post directed to U. S. constitutional lawyers was posted on [Balkenization](#) on December 4, 2018.

